Strategic Evolution of ESE Data Systems

SEEDS Formulation Approach
February 5, 2002

Dr. Stephen Wharton

SEEDS Formulation Manager

301-614-5350

wharton@eosdata.gsfc.nasa.gov

Presentation Outline

- **SEEDS Formulation Approach:**
  - Formulation Scope
  - Linkage to NewDISS Strategy Document
  - Post-Formulation Notional Scope
  - Formulation Objective
  - Formulation Studies
  - Coordination/Communication Efforts to Date
  - Anticipated Results from Formulation
  - SEEDS Phases
- **Background Material**
  - Lessons Learned Questions to Federation

Formulation Scope

- **Charter: Establish a strategy for the evolution of the Earth Science Enterprise network of data systems and service providers to support the Enterprise science objectives in the 2002-2010 timeframe.**
- **The Formulation Team is following the recommendation of the strategy document that the way to foster evolution is to establish a "framework" of interfaces, protocols, and "levels of service" standards and requirements:**
  - Enable ESE system and service providers to implement their own independent data systems and services.
  - Maintain sufficient system wide interoperability to facilitate discovery and interchange of data, information, and services.
- **SEEDS will not be responsible for managing contract support, implementation, operation, or determining the number, type, or assignment of future ESE data centers and service providers (nor allocation of mission data responsibilities).**
  - This authority will be in the hands of the NASA Programs, Projects, PI's, or science teams selected to develop and operate new Enterprise missions, research, and applications projects.
  - Decisions regarding the priorities and allocations for funding ESE data systems and services will be made external to SEEDS.
- **We envision a SEEDS Office that will be responsible for maintaining and evolving the standards and requirements needed for SEEDS success and for monitoring and reporting on the collective performance of the data system elements towards meeting the Enterprise objectives.**

Linkage to NewDISS Strategy Document

- **We summarize the traceability of SEEDS formulation to the NewDISS Strategy Document in terms of three subject areas:**
  - Unifying Framework - How to introduce greater flexibility and responsiveness into the standards, processes, and infrastructure used to support the generation of science data products from NASA science missions?
    - This question is the principal focus of the current formulation effort.
  - Program/Project Management - How to establish a more broadly based network of science and applications providers of products, system capabilities, and services to address the ESE strategic science objectives?
    - This question is outside of the scope of the formulation effort as currently defined.
    - SEEDS will not be responsible for managing contract support, implementation, operation, or determining the number, type, or assignment of future ESE data centers and service providers.
    - This authority will be in the hands of the NASA Programs, Projects, PI's, or science teams selected to develop and operate new Enterprise missions, research, and applications projects.
    - SEEDS does expect to play a supporting role by assisting NASA HQ in developing specifications for AO/NRA's and in evaluating the proposals received.
  - Technology Infusion - How to establish an integrating framework of data management standards, processes, and infrastructure that leverages information technologies that exist or are being developed by other agencies/commercial sector?
    - This question will be addressed by working with the Earth Science Technology Office at GSFC to evaluate future data management technology infusion needs/opportunities and to identify budget requirements.
- **A separate presentation provides a more in-depth linkage to the NewDISS Strategy Document – Recommendations, Lessons Learned, Principles for NewDISS, and Core Standards and Practices.**

Post-Formulation Notional Scope

- **ESE Enterprise:**
  - Establishes overall data management funding profile.
  - Sets budget, priorities, and requirements for data systems and services in support of missions, science measurements, applications, and education.
  - Issues calls for proposals, evaluates proposals, and awards funding to selected teams.
  - Selects appropriate NASA Program(s)/Project(s) to oversee development and operation.
- **SEEDS "Office":**
  - With community, establishes and evolves framework of standards, levels of service, and interfaces.
  - Supports Enterprise data management resource planning.
  - Prepares specifications for ESE calls for proposals for new science missions and services.
  - Sponsors reviews of proposals and compiles performance metrics of data services providers.
- **NASA Implementing Program/Project:**
  - Establish requirements, plans, schedules, deliverables with the selected implementing team(s).
  - Manages funding, monitors progress, and ensures compliance with standards and requirements.
  - Provides data management infrastructure support as appropriate.
  - Ensures that systems and products meet NASA security and survivability requirements.
- **Implementing Team:**
  - Defines science and/or applications requirements for products and services.
  - Designs and implements products and services.
  - Develops approach, selects partners, and carries out design, implementation, and operation.

Formulation Objective

- **Recommend a unifying framework, programmatic guidelines, and processes to enable evolution towards a future network of ESE data systems and providers that:**
  - Sustains and leverages ESE's existing data system operability as appropriate.
  - Enables development and evolution of heterogeneous systems and services that gives systems and service providers appropriate local control over data system design, implementation, and operation.
  - Leverages existing community standards to the greatest extent possible.
  - Engages community in the ongoing definition/evolution of levels of service, standards, and interfaces for future missions.
  - Leverages competition, technology infusion, and reuse to improve system effectiveness.
  - Ensures that products and services meet norms for utility and accessibility from initial development through long-term archive.
  - Establishes evaluation processes and metrics to monitor collective data service provider performance in meeting the Enterprise objectives and goals.
  - Establishes a linkage between requirements, levels of service, and cost to allow effective resource management and implementation for NASA to carry out its science mission.

Formulation Studies

- **To address the recommendations from the Strategy Document the Formulation Team has established six study team groups:**
  - Cost Estimation and Levels of Service
  - Standards and Interfaces for Near Term Missions
  - Standards and Interfaces for Future Missions
  - Life Cycle Data Management for Long Term Archive
  - Software Reuse and Reference Architecture
  - Metrics Planning and Reporting
- **The study team approach is to:**
  - Establish contract support tasks to provide a coordination and supporting role.
  - Leverage community expertise through consulting arrangements thru support contractors.
  - Solicit ideas for additional participants from members of the ESIS committee.
  - Iteratively:
    - Survey existing practices, capabilities, and lessons learned.
    - Develop a preliminary list of questions.
    - Engage the community in:
      - Clarifying technical areas / questions to be addressed.
      - Identifying science concerns / issues pertinent to the study.
      - Developing and reviewing options to address concerns / questions.
      - Developing and refining SEEDS recommendations.
    - Revise/refine survey, questions, concerns, recommendations in response to community feedback.
  - Disseminate findings, status, and draft results via workshops and web page.

Coordination/Communication Efforts to Date

- **Earth Science Information Partners Federation:**
  - Requested Federation Lessons Learned (May 2001)
  - Presented NewDISS overview at Federation summer meeting
  - Supported Federation's NewDISS Study RFP and will monitor progress and leverage results as they become available.
- **DAACs:**
  - Provided briefing to DAAC Alliance
  - The Alliance has produced a 15 page DAAC response to the lessons learned questions.
  - Provided briefings at DAAC UWG meetings (ORNL, JPL, GSFC, NSIDC)
- **Advisory groups:**
  - Briefed NRC task group on *The Availability and Usefulness of NASA's Space Mission Data*
  - Briefed ESSAAC and ESIS committees
  - Requested that ESIS recommend additional members to strengthen community participation in the study team efforts.
- **Earth Science Technology Office:**
  - Attended ESTO Workshop NASA's Earth Science Technology Conference 2001
  - Planning joint workshop to discuss future technology needs.
  - Discussing technology infusion processes with ESTO

Anticipated Results from Formulation

- **The primary results from the formulation will be in the form of recommendations to the Earth Science Enterprise regarding:**
  - level of service guidelines.
  - standards and interfaces for near-term (2002 – 2004) missions starting development.
  - guidelines for data lifecycle and long term archive planning.
  - process to establish standards and interfaces for future (2004+) missions.
  - Reuse architecture to reduce cost and risk.
  - Metrics to monitor results from data system and service providers.
- **In addition the Formulation Team will:**
  - develop a cost model and benchmarks.
  - Identify future data management technology infusion needs/opportunities.

SEEDS Phases

- **Formulation-Phase I (FY 2002):**
  - Conduct Formulation Studies:
  - Identify future data management technology infusion needs/opportunities.
- **Formulation-Phase II (FY 2003):**
  - Complete cost model and benchmark development.
  - Integrate study results into policy guidelines.
  - Recommend roles and responsibilities for NewDISS Office.
  - Present recommendations and plans to ESIS, ESSAAC, and NRC.
  - Submit recommendations to Enterprise AA for approval.
- **Ideas for Execution Phase (FY 2004):**
  - Establish AO/BAA/CAN mechanism to competitively fund broad community participation in:
    - Supporting the ESE and its projects and user community in the definition and effective use of standards and interfaces for data and information systems.
    - Establishing reuse architecture working groups to move ESE data systems, over the long term, to greater cost effectiveness, responsiveness, flexibility, and openness.
    - Supporting the infusion of advanced technology and the adoption of new standards into currently existing ESE data systems.
  - Provide programmatic guidelines in support of ESE data system related AO's and NRA's and assist in the technical and cost analysis of proposals submitted.
  - Compile and report metrics that monitor collective data service provider performance in meeting the Enterprise objectives and goals.

Background Slides

Lessons Learned Questions to Federation

- **1. What has the Federation / NewDISS Cluster learned in terms of the analysis, development, and adoption of standards necessary to facilitate the exchange of data, information, and services among a network of distributed, heterogeneous data and information service providers? More specifically:**
  - What barriers has the Federation encountered that limit the access, dissemination, and utilization of Earth Science data?
  - What is the relative importance of data format standards, metadata standards, search and access services, interface standards, data subsetting and translation services to address barriers encountered by the science community?
  - What is the relative importance of data format standards, metadata standards, search and access services, interface standards, data subsetting and translation services to address barriers encountered by the applications community?
  - Given the Federation's experience to date: a) what are the strengths and shortcomings of the current suite of standards in use within the science and applications communities; b) how would the Federation now "qualify" a standard as having been accepted / adopted for use by the community; and c) how would the Federation distinguish the scope of "core" versus "community" standards?

Lessons Learned Questions to Federation - 2

- **2. What is the appropriate balance of priorities and funding between "top-down" efforts such as the development of standards and interfaces versus "bottom-up" formulation of prototypes and clusters? This question is at the root of the strategic issue of how to determine what should be centrally managed and what should not. More specifically:**
  - What activity does the Federation consider to be more vital and effective – standards or prototyping? On what basis should the proportion of funding for standards development versus prototype development be determined?
  - What is the appropriate balance of funding for Federation activities? What portion of science and application "value-chains" is appropriate for funding by Enterprise science and applications initiatives and what portion is appropriate for funding by the NewDISS Program?
- **3. Given the Federation's experience with governance, would the Federation recommend a differing governance approach in the future? Specifically:**
  - Would the Federation desire more or less involvement from NASA? (Is the Federation of one mind on this topic, or are there substantial groups with differing views and with differing reasons?)
  - If the Federation would desire more involvement from NASA, how would it characterize that involvement?
  - Does the Federation see any need to re-organize itself for alignment with changing ESE plans? If it does see a need to reorganize, how would it do that?
  - How should other agencies engaged in Earth System Science data and information be built into NewDISS?
  - Has the current competitive model, with its associated intellectual property issues, been a help or a hindrance to progress in the federation? Would other models, such as Open Source after initial competitive selection, be more useful?